UNITED STATES PATENT APPLICATION

FOR

A METHOD AND APPARATUS FOR DISCOVERING CLIENT PROXIMITY USING
IN-LINE TRANSLATIONS

Inventor:

Peter A. Tenereillo
William LeBlanc
James A. Jordan
Richard A. Howes, Jr.

Prepared by:

A METHOD AND APPARATUS FOR DISCOVERING CLIENT PROXIMITY USING IN-LINE TRANSLATIONS

This application claims the benefit of priority on U.S. Provisional Patent Application 60/318,426 filed September 10, 5 2001.

Field

The invention generally relates to the field of communications. More particularly, the invention relates to the management of traffic over a network.

10 General Background

As usage of the Internet has increased, various web sites have responded by adding such features as redundancy and load balancing. Accordingly, it has become important to direct a client to the geographically closest and least busy web site. 15 The practice of dispersing data from a web site closer to the client is referred to as "client proximity" and has been widely accepted.

In an attempt to direct clients to the geographically closest and least busy web site, a number of techniques have been 20 implemented. Current technologies that attempt to provide a measure of geographic distribution of Web requests rely primarily on Domain Name System (DNS) techniques; namely, techniques normally performed by a Domain Name System (DNS) server alone or in combination with other logic.

25 One prior technique is Round-Robin Domain Name System (DNS). This technique involves entering multiple IP addresses to represent a single DNS hostname. As clients resolve the

1

hostname, DNS responds by cycling through the multiple listed IP addresses.

Another technique involves computation of routing metrics by the DNS server and these metrics being used to determine how far away a client is from various web sites forming the global domain. This allows the DNS server to answer a DNS request with a web site address associated with a local domain considered to be closest to the client. The primary DNS server can determine the distance from the client to each web site by counting network hops.

Another technique involves the use of a DNS server in conjunction with routers to approximate network distance to a web site from the requestor. This technique is achieved through the announcement of a single IP address or a single set of IP addresses throughout the Internet resolving to a single hostname.

Yet another technique involves geographically distributed DNS servers that provide differing IP addresses on a per server basis. This technique is achieved through the announcement of a single IP address for authoritative DNS servers, which when queried may each provide a different response specifying the nearest web site.

While these DNS techniques provide some load sharing capabilities, they are inherently problematic because they are difficult and resource-intensive to resolve. In addition, the DNS solutions are incapable of being content aware and are, at best, useful in assisting a more robust approach by initially guiding a client to a web resource.

There are now attempts to develop a client proximity process using personal content directors (i.e. site selectors) in which

2

clients are directed to "chronometrically optimal" locations, namely locations that provide the best overall response time, taking into account all factors including geographical distance, network topology (latency), server response time and the like.

5   Exemplary client proximity processes are described in commonly-owned, U.S. Patent Application No. 09/728,305 (filed November 30, 2000), entitled "Method and Apparatus For Discovering Client Proximity". For instance, during "Image Insert" mode, the client proximity process avoids added latency associated with other

10   processes, but does not avoid added network traffic.

3

SUMMARY

One embodiment of the invention relates to a system and method for determining a chronometrically optimal location for a client based on proximity measurements on established connections that are a result of requests for actual content. Instead of adding transparent images or other types of objects to a downloaded web page, the links are altered to enable PCDs to compute return trip time (RTT) values and provide such values to the synchronizing PCD for subsequent selection of the optimal site.

4

BRIEF DESCRIPTION OF THE DRAWINGS

The features and advantages of the invention will become apparent from the following detailed description of the invention in which:

5   Figure 1 is an exemplary embodiment of a communication system utilizing personal content directors (PCDs) according to the invention.

Figure 2 is an exemplary embodiment of a personal content director (PCD) of Figure 1.

10   Figure 3 is an exemplary flowchart of the operations conducted by a selected, synchronizing personal content director.

Figure 4 is an exemplary embodiment of a HTML web page with a plurality of objects displayed thereon according to contents extracted by the synchronizing PCD.

15   Figure 5 is an exemplary embodiment of a translation operation conducted by the synchronizing PCD.

Figure 6 is an exemplary embodiment of an HTML web page with translated objects as received by the client from the synchronizing PCD.

20   Figure 7 is an exemplary flowchart of the operations conducted by the client upon receiving the translated HTML web page from the synchronizing PCD as further shown in Figure 1.

Figure 8 is an exemplary embodiment of a communication system utilizing a load-balancing switching device to resolve a
25   request into a virtual IP address (VIP) targeted for one of multiple personal content directors (PCDs).

Figure 9 is an exemplary embodiment of an HTML web page with tagged objects as received by the client from the synchronizing PCD.

6

In general, one embodiment of the invention relates to a system and method for determining a chronometrically optimal location for a client based on proximity measurements on established connections that are a result of requests for actual content. No transparent Graphic Interface Format (GIF) images or other types of objects are added to a downloaded web page.

Certain details are set forth below in order to provide a thorough understanding of the invention, albeit the invention may be practiced through many embodiments other that those illustrated. Well-known logic and operations are not set forth in detail in order to avoid unnecessarily obscuring the invention.

In the following description, certain terminology is used to describe certain features of the invention. For example, a "personal content director" or "PCD" is a computing device that is adapted to a network (e.g., the Internet) in order to optimize performance of domains hosted on geographically distributed, mirrored web sites. Normally, a PCD comprises internal logic, namely hardware, firmware, software module(s) or any combination thereof, having an architecture generally equivalent to a computer (e.g., server, desktop, laptop, hand-held, mainframe, or workstation), set-top box, or a network switching device such as a router, bridge or brouter.

A "client" is a computing device that executes Web Browser software to communicate over a network in order to download information from a web site. Such information may include HyperText Markup Language (HTML) web pages.

A "software module" is a series of instructions that, when executed, performs a certain function. Examples of a software module include an operating system, an application, an applet, a program or even a routine. One or more software modules may be stored in a machine-readable medium, which includes but is not limited to an electronic circuit, a semiconductor memory device, a read only memory (ROM), a flash memory, a type of erasable programmable ROM (EPROM or EEPROM), a floppy diskette, a compact disk, an optical disk, a hard disk, or the like.

Referring to Figure 1, an exemplary embodiment of a communication system utilizing personal content directors (PCDs) is shown. Herein, system (100) comprises a client 110 in communication with a plurality of geographically dispersed, fully replicated web sites 120_1-120_N over a network 130, where "N" is equal to three for this embodiment. Of course, other embodiments may exist with more than three replicated web sites ($N \geq 4$).

As shown, web site 120_1 includes a switching device 140_1 for routing information, a personal content director (PCDs) 150_1 and a server farm 160_1. A "server farm" is generally defined as a server or multiple servers operating in a collective manner. Similarly, web sites 120_2 and 120_3 include switching devices 140_2, 140_3; PCDs 150_2, 150_3 and server farms 160_2, 160_3, respectively.

During the client proximity selection process, each PCD 150_1-150_3 is adapted to operate as a proxy server to communicate requests by client 110 to one or more servers 160_1-160_3. During normal operations, however, the servers 160_1-160_3 are in communication with the client 110 over network 130 as shown by routing path 161_1-161_3. Moreover, each PCD 150_1-150_3 is further configured to communicate with each other and client 110 using appropriate Transmission Control Protocol (TCP) over network 130.

8

For this embodiment, network 130 is any type of wide area network (WAN) such as the Internet. Of course, network 130 may be configured as a type of local area network (LAN) while still maintaining the spirit and scope of the invention.

Although not shown, system 100 further includes an authoritative Domain Name Server (DNS) that resolves (translates) an alphanumeric, web site domain name into addresses recognized by the PCDs 150_1-150_3. For the system 100, multiple distributed web sites 120_1-120_3 appear to be a single domain (referred to as a "global domain"). A global domain may be separated into multiple "real" web sites, each referred to herein as "local domains" and uniquely registered in DNS with a unique uniform resource locator (URL). To participate in the client proximity selection process, a PCD needs to be configured as a member of that local domain. Each PCD can have more than one global domain; a mirror web site can support more than one local domain; and a PCD can be a member of more than one local domain.

For example, in one embodiment, the web site domain "www.nortelnetworks.com" may be resolved into one of a plurality of IP addresses that is associated with a participating PCD 150_1-150_3 of web sites 120_1-120_3. These IP addresses are entered into DNS as the addresses for the web site domain as shown in Table 1. Such selection as to which IP address may be controlled by a round-robin determination for example.

DNS Name	Participant PCDs
www.nortelnetworks.com	PCD A, B, C
wwwa.nortelnetworks.com	1.1.1.1 (PCD A)

9

wwwb.nortelnetworks.com	2.2.2.2 (PCD B)
wwwc.nortelnetworks.com	3.3.3.3 (PCD C)

TABLE 1

Referring now to Figure 2, an exemplary embodiment of a personal content director (PCD) 150_1 of Figure 1 is shown. PCD 150_1 comprises a network interface 200, a processor 210 and memory 220. Network interface 200 supports multiple TCP communication streams 230_1-230_M. One TCP stream 230_1 may be a Hypertext Transfer Protocol (HTTP) GET request in accordance with a client proximity selection process described below. Another TCP stream 230_M may be information regarding response time statistics transmitted in accordance with a message digest based (MD5) authenticated protocol running on TCP as set forth in commonly-owned, U.S. Patent Application No. 09/728,305 incorporated herewith by reference.

Processor 210 controls the retrieval of content from server(s) at the web site 120_1 of Figure 1. Examples of processor 210 may include, but is not limited or restricted to a microprocessor, digital signal processor, application specific integrated circuit (ASIC), microcontroller, and the like.

Memory 220 is loaded with (i) software modules executable by processor 210 to support client proximity selection process operations of web site 120_1 and (ii) a client network cache (CNC) 240. CNC 240 is configured to store representations of client network addresses associated with client/local domain responses. Upon receiving an initial client request (e.g., HTTP GET request), PCD 150_1 acts as a synchronizing PCD.

10

Upon finding an entry 250 in its CNC 240, synchronizing PCD 150_1 would direct client 110 of Figure 1 to the associated web site, thus avoiding data overhead for client proximity selection. CNC 240 can correlate user data based on DNS name of the client or IP address and compare this data to allow a client that has yet not established a connection with synchronizing PCD 150_1 to benefit from previous client connections. If no entry is found, however, synchronizing PCD 150_1 may enter into a particular mode of operation for the client proximity selection process referred to as an in-line translation (ILX) mode, described below.

To support in-line translation (ILX) mode, the IP address of each local domain and its subsequently computed return trip time (RTT) measurements are stored in one or the entries 250 of CNC 240. Each entry may also contain a date-stamp, which allows for deletion of entries that exceed administrator-specified time limits.

A.   First Embodiment

Referring now to Figure 3, an exemplary flowchart of the operations conducted by a selected, synchronizing personal content director (PCD) (or software executed therein) is shown. Initially, during the client proximity process, a web page request (e.g., HTTP GET request) is received by a PCD, now referred to as the synchronizing PCD (block 300). As briefly stated above, the DNS server chooses the synchronizing PCD and forwards the web page request thereto.

Upon receiving the request, the synchronizing PCD proxies server(s) of its Web site for requested content forming the Web page (block 310). For instance, the requested content may form a home page (default.html). The type of content may include text

11

and images as shown in Figure 4. For instance, the Web page 400 may include three selected references being images 410, 420 and 430 compressed in accordance with any compression function such as Joint Photographic Experts Group (JPEG) for example. Of

5 course, the selected references may be other image types.

Referring to Figures 3, 5 and 6, as described in blocks 320 and 330, synchronizing PCD receives the content from the server(s) and translates selected references to point to its own local domain (wwwa.nortelnetworks.com) as well as other participating

10 local domains (wwwb.nortelnetworks.com and wwwc.nortelnetworks.com). Thereafter, the translated HTML web page 600 of Figure 6 is downloaded to the client (block 340).

For instance, as shown in Figure 5, first selected reference 410 is translated from its relative link " /images/picture1.jpg",

15 returning to a base URL (site A), into an absolute link 510 identified as the following: "http://wwwa.nortelnetworks.com/images/picture1.jpg". Similarly, second selected reference 420 is translated from relative link "/images/picture2.jpg" into an absolute link

20 "http://wwwb.nortelnetworks.com/images/picture2.jpg" (site B) and third selected reference 430 is translated from its relative link "/images/picture3.jpg" into absolute link "http://wwwc.nortelnetworks.com/images/picture3.jpg" (site C).

Referring now to Figure 7, an exemplary flowchart of the

25 operations conducted by the client upon receipt of the translated HTML web page and subsequent operations by the synchronizing PCD is shown. Upon receiving the translated HTML web page (block 700), which now includes absolute links, the client generally follows the links to remote sites and downloads images and other

30 requested files. In addition, the remote PCDs conduct RTT

12

measurements and report these RTT values back to the synchronizing
PCD. For instance, RTT measurements may be accomplished by
computing a time difference between the arrival of the HTTP GET
request and the time of arrival of the final acknowledgement
packet at the PCD. Thereafter, on the next HTTP GET request from
the client, the synchronizing PCD will redirect the client to the
most proximate site.

In particular, for one embodiment, the client makes requests
to local domains for all participating PCDs (block 710). Such
requests may be made to all participating PCDs, being inclusive or
exclusive of the synchronizing PCD. It is exclusive if RTT
measurements were conducted by the synchronizing PCD during
download of the translated HTML web page. For this embodiment,
the requests are made to all PCDs 150_1-150_3 of Figure 1 including
the synchronizing PCD.

Herein, as an exemplary illustration, the client makes a
request to a remote PCD (e.g., PCD 150_2) to download an image
(e.g., a JPEG image "picture2.jpg"). The remote PCD downloads
the image to the client and measures the RTT value, which is
stored within the CNC of the remote PCD (blocks 720 and 730).
This RTT measurement may occur contemporaneously with the
downloading of the image. The RTT value is then reported back to
at least the synchronizing PCD and perhaps to all participating
PCDs (block 740). Such reporting may be made over inter-box
protocol (IBP) based communications being authenticated
communications as described in U.S. Patent Application No.
09/728,305 incorporated herewith by reference. This process is
repeated for all participating PCDs, the order in which these
requests are completed is a design choice (block 750).

13

Thereafter, the participating PCDs for the global domain store the network address of the client. Upon the next GET HTTP request from the client, the client is directed to the optimal site through either DNS or HTTP redirect (block 760).

5   B.   Second Embodiment Using PCD tagging

As shown, different IP addresses are used for the PCDs. However, as an alternative, multiple PCDs 810_1-810_R ($R \geq 1$) may be collectively coupled to a switching device 820 as shown in Figure 8. For this embodiment, the DNS name wwwa.nortelnetworks.com

10   resolves to a virtual IP address (VIP) on load-balancing switching device 820. Herein, the same VIP is used for the RTT measurements. This may be accomplished by placing a URL tag (e.g., an alphanumeric string) into the tagged (URL) link in such a manner that parsing rules followed by the load-balancing

15   switching device 820 direct all requests that contain the URL tag to a PCD of the PCDs 810_1-810_R, rather than to the web server, media server, or whatever server from which the content originates.

For example, similar to Figure 6, the translated HTML web

20   page 900 of Figure 9 is downloaded to the client but includes tagged absolute links 910, 920 and 930 (referred to as "tagged links"). Namely, the first reference is translated from its relative link " /images/picture1.jpg", returning to a base URL (site A), into tagged link 910 identified as the following:

25   http://wwwa.nortelnetworks.com/ pcdtag/images/picture1.jpg.

Similarly, the second selected reference is translated from relative link "/images/picture2.jpg" into tagged link "http://wwwb.nortelnetworks.com/pcdtag/images/picture2.jpg" (920,

30   site B) and the third selected reference is translated from

14

relative link "/images/picture3.jpg" into tagged link
"http://wwwc.nortelnetworks.com/pcdtag/images/picture3.jpg" (930,
site C).

Upon receiving the translated HTML web page including tagged
links, for one embodiment, the client generally issues requests
to the local domains. Such requests may be made to switching
devices of all participating PCDs, such as a request
(http://wwwb.nortelnetworks.com/pcdtag/images/
picture2.jpg) to a remote PCD for downloading an image (e.g., a
JPEG image "picture2.jpg"). The switching device 820 operates to
detect the URL tag "pcdtag" and directs the request to one of the
PCDs (e.g., PCD 810_1). The PCD 810_1 strips the URL tag and
proxies the request to the server farm 830 associated with the
PCD 810_1. The PCD 810_1 downloads the retrieved image to the
client and measures the RTT value, which is stored within the CNC
of the PCD 810_1 and the process continues as described above.